

Doublon-holon binding, Mott transition, and fractionalized antiferromagnet in the Hubbard model

Sen Zhou^{1,2}, Yupeng Wang³, and Ziqiang Wang²

¹ *State Key Laboratory of Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China*

² *Department of Physics, Boston College, Chestnut Hill, MA 02467, USA and*

³ *The Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China*

(Dated: February 25, 2013)

We argue that the binding between doubly occupied (doublon) and empty (holon) sites governs the incoherent excitations and plays a key role in the Mott transition in strongly correlated Mott-Hubbard systems. We construct a new saddle point solution with doublon-holon binding in the Kotliar-Ruckenstein slave-boson functional integral formulation of the Hubbard model. On the half-filled honeycomb lattice, the ground state exhibits a continuous transition from the semimetal to an antiferromagnetic ordered Slater insulator with coherent quasiparticles at $U_{c1} \simeq 3.4t$, followed by a Mott transition into an electron-fractionalized AF* phase without coherent excitations at $U_{c2} \simeq 5.7t$. We show that doublon-holon binding unites the three important ideas of strong correlation: the coherent quasiparticles, the incoherent Hubbard bands, and the deconfined Mott insulator.

PACS numbers: 71.10.-w, 71.27.+a, 71.10.Fd, 73.30.+h

The fundamental theoretical challenge of the strong correlation problem is the description of both the low energy coherent quasiparticles (QPs) and the higher energy incoherent excitations, as well as the spectral weight transfer from coherent to incoherent excitations with increasing correlation strength. Two very important ideas, the emergence of two broad incoherent features known as the Hubbard bands and the existence of renormalized QPs with a Luttinger Fermi surface (FS) were advanced by Hubbard [1], and Brinkman and Rice [2], respectively. Unfortunately, the Hubbard equation of motion scheme that produces the incoherent spectrum fails to produce QPs correctly and violates Luttinger's theorem [3]; whereas the approaches based on the Brinkman-Rice-Gutzwiller wave functions [4] capture a Luttinger FS of QPs, but find serious difficulties in constructing variational excited states to account for the incoherent excitations. Faced with this enigma, numerical approaches such as exact diagonalization, quantum Monte Carlo (QMC), and the dynamical mean field theory [5] have played a key role in recent years. In this paper, we develop new analytical insights, construct a unified theory for both coherent and incoherent excitations and the Mott transition, and study the Hubbard model on the honeycomb lattice in view of the recent debate over the emergence of a gapped spin liquid (SL) phase [6–10].

For simplicity, we consider the half-filled single-band Hubbard model. The local Hilbert space consists of empty (holon), doubly occupied (doublon), and singly occupied sites. The Brinkman-Rice-Gutzwiller approach amounts to a metallic state where the holon, denoted as a boson e , and the doublon, as d , condense fully with macroscopic phase coherence, as obtained by the Gutzwiller approximation [11] or the slave boson mean-field theory (SBMFT) [12]. The metal-insulator transi-

tion is thus forced to follow a route where the *density* of doublons and holons vanish together with the condensates: $n_d = n_e = \langle d \rangle = \langle e \rangle = 0$ such that there is *exactly* one electron per site. As a result, single-particle motion, coherent or incoherent, is completely prohibited. This so-called Brinkman-Rice (BR) transition is different from the Mott transition induced by the complete transfer of the coherent QP weight into the incoherent background, *i.e.* the depletion of the condensate while keeping the doublon/holon (D/H) *density* an analytic function in the correlation strength U .

We will show that the crucial physics uniting the disparate ideas of Hubbard and Brinkman-Rice is the binding between doublon and holon: $\langle d_i e_j \rangle \neq 0$. In the Mott insulator at large U , although the D/H condensate vanishes ($\langle d \rangle = \langle e \rangle = 0$) together with coherent QP, the D/H *density* remains nonzero ($n_d = n_e \neq 0$). The incoherent motion of the QP is thus possible by breaking the doublon-holon (D-H) pairs, giving rise to the higher-energy incoherent excitations. With decreasing U , the D/H density increases and the D-H binding energy decreases. At a critical U_c , the D-H excitation gap closes and the D/H single-particle condensate develops, marking the Mott transition. On the metallic side of the Mott transition, D-H binding continues to play an important role since an added electron can propagate either as a coherent QP via the D/H condensate or incoherently via the unbinding of the D-H pairs.

The idea that D-H binding plays a role in Mott-Hubbard systems was introduced by Kaplan, Horsch, and Fulde [13] and studied in the context of improved variational Gutzwiller wave functions [14, 15]. It has been difficult to advance because of the difficulty in the calculations and particularly in constructing the appropriate wave functions for excitations. We will show that the

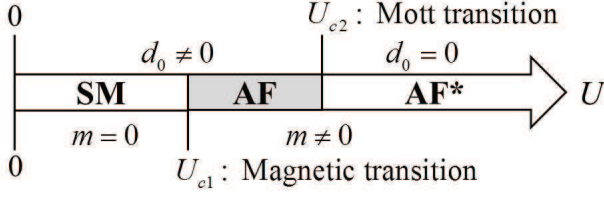


FIG. 1: Schematic phase diagram obtained for the half-filled honeycomb lattice Hubbard model with D-H binding.

physical picture discussed above can be realized in the slave-boson functional integral formulation of the Hubbard model introduced by Kotliar and Ruckenstein (KR) [12] by constructing new saddle point solutions that include the D-H binding. This approach also offers a treatment of the magnetism at half-filling that compares well with QMC simulations [16], and has the added advantage of allowing the study of excitations and finite temperature properties [17].

As a concrete example, we studied the D-H binding in the half-filled Hubbard model on the honeycomb lattice at zero temperature and obtained the phase diagram shown schematically in Fig. 1. A continuous transition from the semimetal (SM) to an antiferromagnetic (AF) ordered insulator takes place at a critical $U_{c1} \simeq 3.4t$, suggesting that the gapped SL phase proposed by Meng et. al. [6] may correspond to an AF ordered phase in the thermodynamic limit. Sorella et. al. [10] recently extended the QMC and the finite size scaling analysis to much larger system sizes and discovered that the signature of the gapped SL disappears and is replaced by that of a continuous SM to AF order transition at $U \simeq 3.8t$, in qualitative agreement with our results. Remarkably, we found a second quantum phase transition at a critical $U_{c2} \simeq 5.7t$ beyond which the D/H single-particle condensate vanishes ($d_0 = 0$) amid a finite density of doublons bound to holons. For $U > U_{c2}$, a new AF phase, termed as AF* in Fig. 1, emerges where the electrons are fractionalized and there are no coherent QP excitations.

In KR's theory, the electron operator is written as $c_{i\sigma} = \hat{L}_{i\sigma}(e_i^\dagger p_{i\sigma} + p_{i\bar{\sigma}}^\dagger d_i) \hat{R}_{i\sigma} f_{i\sigma}$, where the boson operators describe the holon (e_i), doublon (d_i), and singly-occupied sites ($p_{i\sigma}$), and $f_{i\sigma}$ is a fermion operator. The operators \hat{L}_σ and \hat{R}_σ are diagonal with unit eigenvalues in the (empty, $\bar{\sigma}$) and the (σ , doubly-occupied) subspaces, respectively, $\hat{L}_{i\sigma} = (1 - d_i^\dagger d_i - p_{i\sigma}^\dagger p_{i\sigma})^\alpha$ and $\hat{R}_{i\sigma} = (1 - e_i^\dagger e_i - p_{i\bar{\sigma}}^\dagger p_{i\bar{\sigma}})^\beta$. KR found that for $\alpha = \beta = -1/2$, the saddle point solution recovers the Gutzwiller results. The Hubbard model Hamiltonian is thus given by [12],

$$\hat{H} = -t \sum_{\langle i,j \rangle, \sigma} (\hat{z}_{i\sigma}^\dagger \hat{z}_{j\sigma} f_{i\sigma}^\dagger f_{j\sigma} + h.c.) + U \sum_i d_i^\dagger d_i + i \sum_i \lambda_i \hat{Q}_i + i \sum_{i,\sigma} \lambda_{i\sigma} \hat{Q}_{i\sigma} - \mu \sum_{i\sigma} f_{i\sigma}^\dagger f_{i\sigma}, \quad (1)$$

where $\hat{z}_{i\sigma} = \hat{L}_{i\sigma}(e_i^\dagger p_{i\sigma} + p_{i\bar{\sigma}}^\dagger d_i) \hat{R}_{i\sigma}$ and the Lagrange

multipliers λ_i and $\lambda_{i\sigma}$ enforces the local constraints: $\hat{Q}_i = e_i^\dagger e_i + \sum_\sigma p_{i\sigma}^\dagger p_{i\sigma} + d_i^\dagger d_i - 1 = 0$ and $\hat{Q}_{i\sigma} = f_{i\sigma}^\dagger f_{i\sigma} - p_{i\sigma}^\dagger p_{i\sigma} - d_i^\dagger d_i = 0$. The partition function is a functional integral over the quantum fields [22]: $Z = \int \mathcal{D}[f, e, p, d, \lambda] \exp(-\int_0^\beta \mathcal{L} d\tau)$, where the Lagrangian $\mathcal{L} = \sum_i (e_i^\dagger \partial_\tau e_i + d_i^\dagger \partial_\tau d_i) + \sum_{i,\sigma} (p_{i\sigma}^\dagger \partial_\tau p_{i\sigma} + f_{i\sigma}^\dagger \partial_\tau f_{i\sigma}) + \hat{H}$. The KR saddle point corresponds to condensing all boson fields uniformly in Eq. (1). The results on the honeycomb lattice [23] can be summarized: Restricting to the PM phase, the doublon density d_0^2 decreases linearly from 1/4 at $U = 0$ and vanishes at the BR metal-insulator transition $U_{BR} \simeq 12.6t$. When magnetism is allowed, a SM to an AF insulator transition arises at $U_m \simeq 3.1t$. The D/H condensate remains nonzero for any finite U and the AF phase is a coherent Slater insulator [24].

To construct the new saddle point solution with D-H binding, we need to include the dynamics of the bosons. Introducing the operators for the D-H pairing $\hat{\Delta}_{ij} = d_i e_j$, and the D/H hopping $\hat{\chi}_{ij}^d = d_i^\dagger d_j$, $\hat{\chi}_{ij}^e = e_i^\dagger e_j$ on the nearest-neighbor bond, as well as the density operators $\hat{n}_i^d = d_i^\dagger d_i$, $\hat{n}_i^e = e_i^\dagger e_i$ on each site, we can write in Eq. (1): $\hat{z}_{i\sigma}^\dagger \hat{z}_{j\sigma} = \hat{g}_{ij}^\sigma [(\hat{\chi}_{ij}^e)^\dagger p_{i\sigma}^\dagger p_{j\sigma} + \hat{\chi}_{ij}^d p_{i\bar{\sigma}} p_{j\bar{\sigma}}^\dagger + \hat{\Delta}_{ji} p_{i\sigma}^\dagger p_{j\bar{\sigma}}^\dagger + \hat{\Delta}_{ij} p_{i\bar{\sigma}} p_{j\sigma}]$, where $\hat{g}_{ij}^\sigma = (\hat{Y}_{ij}^\sigma \hat{Y}_{ji}^\sigma)^{-1/2}$, $\hat{Y}_{ij}^\sigma = (1 - p_{i\bar{\sigma}}^\dagger p_{i\bar{\sigma}})(1 - p_{j\sigma}^\dagger p_{j\sigma}) - \hat{n}_i^e(1 - p_{j\sigma}^\dagger p_{j\sigma}) - \hat{n}_j^d(1 - p_{i\bar{\sigma}}^\dagger p_{i\bar{\sigma}}) + |\hat{\Delta}_{ji}|^2$. These inter-site correlations can be treated explicitly by writing the partition function as a path integral over the corresponding fields and taking the saddle-point solution [24]. The results can be understood intuitively as minimizing the variational Hamiltonian obtained by substituting the above expression for $\hat{z}_{i\sigma}^\dagger \hat{z}_{j\sigma}$ into Eq. (1) with the operators replaced by their expectation values, and adding the corresponding variational terms:

$$\hat{H}_{DH} = -t \sum_{\langle i,j \rangle, \sigma} (z_{i\sigma}^\dagger z_{j\sigma} f_{i\sigma}^\dagger f_{j\sigma} + h.c.) + U \sum_i d_i^\dagger d_i + i \sum_i \lambda_i \hat{Q}_i + i \sum_{i,\sigma} \lambda_{i\sigma} \hat{Q}_{i\sigma} - \mu \sum_{i\sigma} f_{i\sigma}^\dagger f_{i\sigma} - i \sum_{\langle i,j \rangle} \left[\chi_{ij}^{d,v} (d_i^\dagger d_j - \chi_{ij}^d) + \chi_{ij}^{e,v} (e_i^\dagger e_j - \chi_{ij}^e) + \Delta_{ij}^v (d_i e_j - \Delta_{ij}) + \Delta_{ji}^v (e_i d_j - \Delta_{ji}) + h.c. \right] + i \sum_i \left[\epsilon_i^{d,v} (d_i^\dagger d_i - n_i^d) + \epsilon_i^{e,v} (e_i^\dagger e_i - n_i^e) \right], \quad (2)$$

where Δ_{ij}^v , $\chi_{ij}^{d,v}$, $\chi_{ij}^{e,v}$, $\epsilon_i^{d,v}$, and $\epsilon_i^{e,v}$ are self-consistently determined variational parameters. Note that the $p_{i\sigma}$ bosons are treated as condensed fields for simplicity since their density (single occupation) is large [25].

We next discuss the saddle point solution of Eq. (2) on the honeycomb lattice with $2N$ sites. Symmetry requires the expectation values $\Delta_{ij} = \Delta_d$, $\chi_{ij}^d = \chi_{ij}^e = \chi_d$, $n_i^d = n_i^e = n_d$, and correspondingly, $i\Delta_{ij}^v = \Delta_d^v$, $i\chi_{ij}^{d,v} = i\chi_{ij}^{e,v} = \chi_d^v$, $i\epsilon_i^{d,v} = i\epsilon_i^{e,v} = \epsilon_d^v$. Moreover, $i\lambda_i = \lambda$, $i\lambda_{A\sigma} =$

$i\lambda_{B\bar{\sigma}} = \lambda_{\sigma}$, and $p_{A\sigma}^{(\dagger)} = p_{B\bar{\sigma}}^{(\dagger)} = p_{0\sigma}$, where A and B denote the two sublattices. The normalization factor $g_{ij}^{\sigma} = g = \prod_{\sigma} Y_{\sigma}^{-1/2}$, with $Y_{\sigma} = 1 - 2n_d - 2p_{0\sigma}^2 + 2p_{0\sigma}^2 n_d + p_{0\sigma}^4 + \Delta_d^2$. As shown later, this expression recovers the noninteracting limit at $U = 0$. The Hamiltonian thus simplifies to $\hat{H}_{DH} = \hat{H}_f + \hat{H}_d + 12N(\chi_d^{\vee}\chi_d + \Delta_d^{\vee}\Delta_d) - 2N(\epsilon_d^{\vee} + 2\lambda) - 4N\epsilon_d^{\vee}n_d + 2N\sum_{\sigma}(\lambda - \lambda_{\sigma})p_{0\sigma}^2$. The fermion spectrum is given by

$$\hat{H}_f = \sum_{\mathbf{k},\sigma} \begin{bmatrix} f_{A\mathbf{k}\sigma} \\ f_{B\mathbf{k}\sigma} \end{bmatrix}^{\dagger} \begin{bmatrix} \lambda_{\sigma} - \mu & -\chi_f^{\vee}\eta_{\mathbf{k}} \\ -\chi_f^{\vee}\eta_{\mathbf{k}}^* & \lambda_{\sigma} - \mu \end{bmatrix} \begin{bmatrix} f_{A\mathbf{k}\sigma} \\ f_{B\mathbf{k}\sigma} \end{bmatrix}, \quad (3)$$

where $\eta_{\mathbf{k}} = \exp(ik_y) + 2\cos(\sqrt{3}k_x/2)\exp(-ik_y/2)$, $\chi_f^{\vee} = tg[2p_{0\uparrow}p_{0\downarrow}\chi_d + (p_{0\uparrow}^2 + p_{0\downarrow}^2)\Delta_d]$, and the chemical potential $\mu = U/2$ by the particle-hole symmetry. The fermion dispersion is thus $E_{\pm}^f(\mathbf{k}) = \pm\sqrt{(\lambda_{\uparrow} - \lambda_{\downarrow})^2/4 + |\chi_f^{\vee}\eta_{\mathbf{k}}|^2}$. A fermion gap $\Xi_f = |\lambda_{\uparrow} - \lambda_{\downarrow}|$ would open in the Dirac spectrum with AF order. Similarly, the charged boson spectrum is determined by $\hat{H}_d = \sum_{\mathbf{k}} \Psi_{\mathbf{k}}^{\dagger} M_{\mathbf{k}} \Psi_{\mathbf{k}}$, where $\Psi_{\mathbf{k}} = [d_{A\mathbf{k}}, d_{B\mathbf{k}}, e_{B\bar{\mathbf{k}}}^{\dagger}, e_{A\bar{\mathbf{k}}}^{\dagger}]^T$ and

$$M_{\mathbf{k}} = \begin{bmatrix} \epsilon_d^{\vee} & -\chi_d^{\vee}\eta_{\mathbf{k}} & -\Delta_d^{\vee}\eta_{\mathbf{k}} & 0 \\ -\chi_d^{\vee}\eta_{\mathbf{k}}^* & \epsilon_d^{\vee} & 0 & -\Delta_d^{\vee}\eta_{\mathbf{k}}^* \\ -\Delta_d^{\vee}\eta_{\mathbf{k}}^* & 0 & \epsilon_d^{\vee} & -\chi_d^{\vee}\eta_{\mathbf{k}}^* \\ 0 & -\Delta_d^{\vee}\eta_{\mathbf{k}} & -\chi_d^{\vee}\eta_{\mathbf{k}} & \epsilon_d^{\vee} \end{bmatrix}. \quad (4)$$

Here $\epsilon_d^{\vee} = \epsilon_d^{\vee} + \lambda$, $\epsilon_d^{\vee} = -3g^2\chi_f\chi_f^{\vee}\sum_{\sigma}(1 - p_{0\sigma}^2)Y_{\sigma}$, and $\chi_d^{\vee} = 2tg p_{0\uparrow}p_{0\downarrow}\chi_f$, $\Delta_d^{\vee} = g\chi_f\sum_{\sigma}(tp_{0\sigma}^2 - g\Delta_d\chi_f^{\vee}Y_{\sigma})$. The boson dispersion is obtained by Bogoliubov transformation: $E_{\pm}^d(\mathbf{k}) = \sqrt{(\epsilon_d^{\vee} \pm |\chi_d^{\vee}\eta_{\mathbf{k}}|)^2 - |\Delta_d^{\vee}\eta_{\mathbf{k}}|^2}$ under the condition $\epsilon_d^{\vee} \geq 3(|\chi_d^{\vee}| + |\Delta_d^{\vee}|)$. The boson energy gap is thus $\Xi_d = 2\sqrt{(\epsilon_d^{\vee} - 3|\chi_d^{\vee}|)^2 - (3|\Delta_d^{\vee}|)^2}$. It is instructive to examine the equations for the D/H density, D/H hopping, and the D-H binding,

$$n_d = d_0^2 + \frac{1}{2N} \sum_{\alpha=\{A,B\}} \sum_{\mathbf{k}} \langle d_{\alpha\mathbf{k}}^{\dagger} d_{\alpha\mathbf{k}} \rangle, \quad (5)$$

$$(\chi_d, \Delta_d) = d_0^2 + \frac{1}{6N} \sum_{\mathbf{k}} \langle \eta_{\mathbf{k}} d_{A\mathbf{k}}^{\dagger} (d_{B\mathbf{k}}, e_{B\bar{\mathbf{k}}}^{\dagger}) + h.c. \rangle. \quad (6)$$

The closing of the boson gap Ξ_d leads to a zero energy mode at $\mathbf{k} = 0$ whose occupation enables the single-boson condensate d_0^2 , and is subsequently taken out of the momentum summations.

The ground state is determined by solving the self-consistency equations derived from minimizing the energy [24]: $\langle \partial \hat{H}_{HD} / \partial x_i \rangle = 0$ and $x = (\Delta_d, \chi_d, n_d, p_{0\sigma}, \lambda_{\sigma}, \lambda)$. Once the saddle point solution is obtained, the spectral function can be obtained from the one-electron Green's function $G_{\alpha\sigma}(\mathbf{k}, \tau) = -\langle T_{\tau} c_{\alpha\mathbf{k}\sigma}(\tau) c_{\alpha\mathbf{k}\sigma}^{\dagger}(0) \rangle$. Since the latter involves convolutions of the d/e boson normal and the anomalous (due to pairing) Green's functions with those of the f_{σ} -fermion, the single-particle

energy gap for the physical electron is the sum of the fermion and boson gaps $\Xi = \Xi_d + \Xi_f$. More importantly, the coherent QP excitations would only emerge with the D/H condensate that recombines the charge and spin degrees of freedom, which can be detected by the coherent peaks in the integrated spectral function (ISF) [24]: $N_{\alpha}(\omega) = -\text{Im} \int_0^{\beta} d\tau e^{i\omega\tau} \sum_{\mathbf{k},\sigma} G_{\alpha\sigma}(\mathbf{k}, \tau)$. Our results are plotted in Figs. 2 and 3. The parameter Δ_d^{\vee} increases with U and the D-H pairing field Δ_d is nonzero for all finite- U (Fig. 2a). At $U = 0$, all doublons and holons are single-particle condensed $d_0^2 = e_0^2 = p_{0\sigma}^2 = 1/4$, recovering the noninteracting limit. The SM phase remains stable at small U . With increasing U , the doublon density decreases. Due to the increase in D-H binding, the D/H condensate d_0 decays faster than in the SBMFT. Let's first restrict the solution to the PM phase with $p_{0\uparrow} = p_{0\downarrow}$. As shown in Fig. 2b, there is a Mott transition at $U_c = 7.3t$, considerably smaller than $12.6t$ for the BR transition. The condensate d_0 vanishes and all doublons are bound with the holons in the Mott insulating phase for $U > U_c$, accompanied by the opening of a charge gap Ξ_d linear in $U - U_c$ (Fig. 2c). The ISF is shown in Fig. 3a. Notice the transfer of the coherent QP weight to the incoherent part with increasing U and the complete suppression of the coherent QPs in favor of two broad incoherent spectra beyond the Mott transition from the two branches of bosonic excitations $E_{\pm}^d(\mathbf{k})$ [24]. Since the f_{σ} -fermion spinon remains gapless, the insulating phase is a gapless SL. We do not see evidence for the proposed gapped SL phase [6].

Next, we allow magnetism and study the interplay between AF order, D-H binding, and the Mott transition in the ground state. Fig. 2b shows the onset of staggered magnetization (m) accompanying the AF transition at a critical $U_{c1} \simeq 3.4t$, which compares well to the continuous SM to AF transition observed in the most recent QMC calculations on large systems [10]. We find that for $U_{c1} < U < U_{c2}$, where $U_{c2} \simeq 5.7t$, although a single-particle gap Ξ_f opens in the fermion sectors (Fig. 2c), the zero energy mode remains stable in the d - e sector and continues to support the D/H condensate. Thus, the spin and charge recombine in this regime and there are coherent excitations corresponding to the sharp QP peaks in the ISF shown in Fig. 3b. This phase is thus an AF Slater insulator. Remarkably, a Mott transition takes place at U_{c2} . For $U > U_{c2}$, a new AF phase, the AF* phase, emerges with the opening of the gap $\Xi_d \propto U - U_{c2}$ in the d - e sector (Fig. 2c), where the D/H condensate vanishes as all doublons are bound to holons. Thus electrons are fractionalized in the AF* phase and there are no coherent excitations, as shown in the broad ISF in Fig. 3b at $U = 6t$. The gap for the physical electron, $\Xi = \Xi_f + \Xi_d$, exhibits a derivative discontinuity at U_{c2} (Fig. 2c) due to the opening of the Mott gap Ξ_d in the AF* phase. Note that the magnetization m is analytic across the AF \rightarrow AF* transition (Fig. 2b), which is a topological confinement-

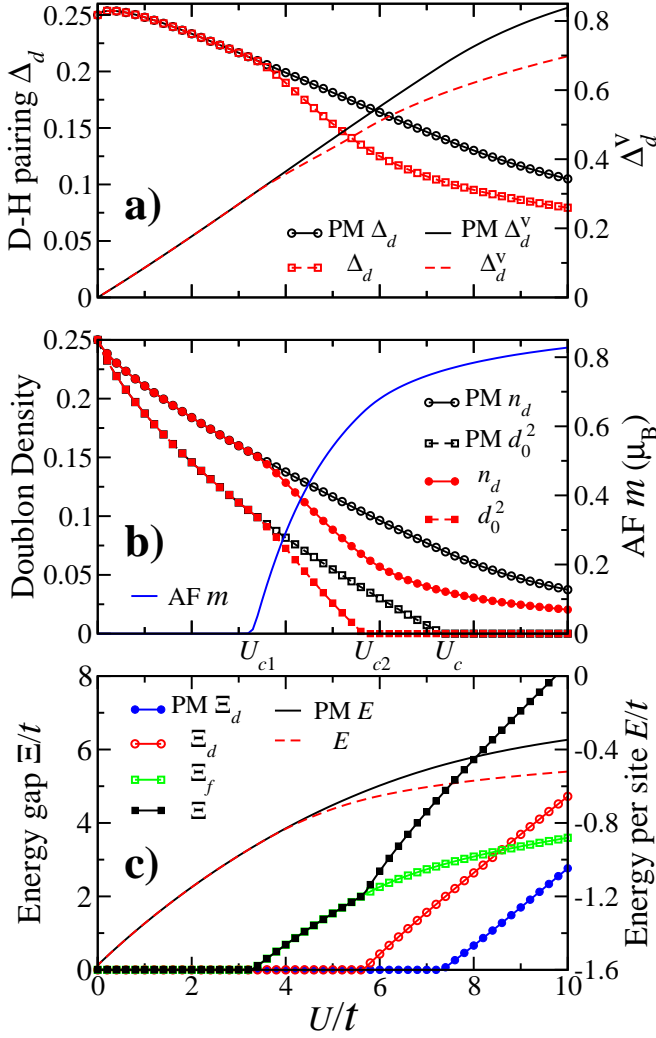


FIG. 2: (color online) Plotted as a function of Hubbard U in (a) are variational parameter Δ_d^v for d - e binding and D-H pairing order parameter Δ_d ; in (b) D/H density n_d , condensate density d_0^2 , and AF staggered magnetization m ; and in (c) ground state energy per site E and energy gaps in the boson Ξ_d and fermion Ξ_f sectors. The corresponding results in the restricted PM phase are also shown.

deconfinement transition associated with the Ising-like global Z_2 symmetry ($d_i \rightarrow -d_i$, $e_i \rightarrow -e_i$) that is broken in the AF phase by the D/H condensate and restored in the AF* phase. Finally, let's examine the saddle point stability with respect to gauge field fluctuations. It is known that KR formulation introduces three $U(1)$ gauge fields [26] since the action is invariant under: $e_i \rightarrow e_i e^{i\theta_i}$, $p_{i\sigma} \rightarrow p_{i\sigma} e^{i\phi_{i\sigma}}$, $d_i \rightarrow d_i e^{-i\theta_i + i\sum_{\sigma} \phi_{i\sigma}}$, $f_{i\sigma} \rightarrow f_{i\sigma} e^{i\theta_i - i\phi_{i\sigma}}$, and $\lambda_i \rightarrow \lambda_i + \dot{\theta}_i$, $\lambda_{i\sigma} \rightarrow \lambda_{i\sigma} + \dot{\theta}_i - \dot{\phi}_{i\sigma}$. The $p_{i\sigma}$ condensate breaks two of the $U(1)$ symmetries and turns the gauge fields associated with $\phi_{i\sigma}$ massive by Anderson-Higgs mechanism. The remaining $U(1)$ symmetry is also broken in the SM and the AF phase by the D/H condensate, making the θ_i -gauge field massive. In the AF* phase, it is the D-H pairing Δ_{ij} that breaks the $U(1)$ symmetry and the θ_i -gauge field remains massive, as does

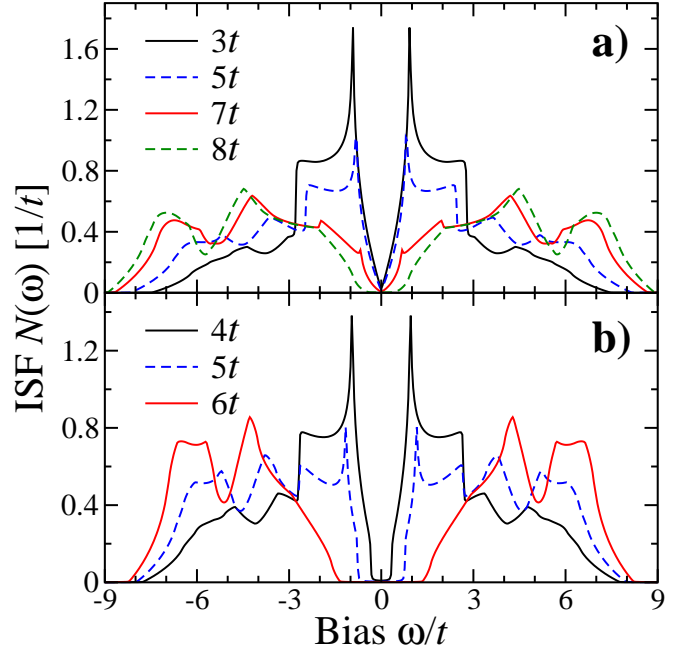


FIG. 3: (color online) Integrated spectral function showing the transfer of coherent QP weight to the incoherent spectra with increasing U listed in the legends. (a) PM phase: Mott transition at $U_c \simeq 7.3t$. (b) Ground state: SM to AF transition at $U_{c1} \simeq 3.4t$ and AF to AF* transition at $U_{c2} \simeq 5.7t$.

its staggered component due to the D/H hopping fields $\chi_{ij}^{d,e}$. The absence of gapless $U(1)$ gauge field fluctuations supports the stability of the obtained phases.

In summary, we have shown that the D-H binding plays an essential role in describing the incoherent excitations and the Mott transition in strongly correlated systems. For the honeycomb lattice Hubbard model, we showed that the SM to AF Slater insulator transition is followed by a Mott transition into a fractionalized AF* phase with increasing U . Interestingly, a different AF* phase of a fractionalized antiferromagnet was proposed in the effective Z_2 gauge theory description of doped Mott insulators in the projected Hilbert space ($U = \infty$) where spinons are paired into a Néel state and doublons are absent [27]. In contrast, the incoherent charge excitations through D-H binding is essential in the AF* phase proposed here, which is more inline with the importance of doublons in describing Motttness emphasized recently [28]. Such an AF* phase on the square lattice may have been observed in the parent compound of the high- T_c cuprates by angle-resolved photoemission [29].

We thank F. Wang and Y. Yu for useful discussions. This work is supported by DOE grants DE-FG02-99ER45747 and de-sc0002554, and the 1000 Talents Plan for Young Researchers of China (SZ).

-
- [1] J. Hubbard, Proc. R. Soc. London A **276**, 238 (1963); **277**, 237 (1964); **281** 401 (1964); **285** 542 (1965); **296**, 82 (1967).
- [2] W. Brinkman and T. M. Rice, Phys. Rev. B **2**, 4302 (1970).
- [3] J. M. Luttinger, Phys. Rev. **121**, 942 (1961).
- [4] M. C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963).
- [5] G. Kotliar, et. al. Rev. Mod. Phys. **78**, 865 (2006); A. Georges, et. al. Rev. Mod. Phys. **68**, 13 (1996).
- [6] Z. Y. Meng, et. al. Nature **464**, 847 (2010).
- [7] W. Wu, et. al. Phys. Rev. B **82**, 245102 (2010).
- [8] A. Liebsch, Phys. Rev. B **83**, 035113 (2011).
- [9] R.-Q. He and Z.-Y. Lu, Phys. Rev. B **86**, 045105 (2012).
- [10] S. Sorella, Y. Otsuka, and S. Yunoki, Sci. Rep. **2**, 992 (2012).
- [11] D. Vollhardt, Rev. Mod. Phys. **56**, 99 (1984).
- [12] G. Kotliar and A.E. Ruckenstein, Phys. Rev. Lett. **57**, 1362 (1986).
- [13] K.A. Kaplan, P. Horsch, and P. Fulde, Phys. Rev. Lett. **49**, 889 (1982).
- [14] H. Yokoyama and H. Shiba, J. phys. Soc. J. **59**, 3669 (1990).
- [15] H. Yokoyama, M. Ogata, and Y. Tanaka, J. Phys. Soc. Jpn. **75**, 114706 (2006).
- [16] L. Lilly, A. Muramatsu, and W. Hanke, Phys. Rev. Lett. **65**, 1379 (1990).
- [17] We do not use the *different* slave-boson construction proposed originally by Barnes [18, 19], since it does not conserve the fermion number and is known to lead to serious unphysical results [20] when charge excitations are present, even for small U [21].
- [18] S. E. Barnes, J. Phys. F **6**, 1375 (1976); *ibid.* **7**, 2637 (1977).
- [19] Z. Zou and P. W. Anderson, Phys. Rev. B **37**, 627 (1988).
- [20] H. Kaga, Phys. Rev. B **46**, 1979 (1992).
- [21] A. Vaezi and X.-G. Wen, arXiv:1010.5744; arXiv:1101.1662; J. Wen, et. al. Phys. Rev. B **84**, 235149 (2011).
- [22] Because of the properties of the projection-operator, the normal ordering of the square roots in the normalization factors is not an issue [12].
- [23] R. Frésard and K Doll, in *The Hubbard Model: Its Physics and Mathematical Physics*, Editors: D. Baeriswyl, et. al., page 385-392, Plenum Press (1995).
- [24] See supplemental section for additional discussions.
- [25] We verified that including the inter-site correlations of the $p_{i\sigma}$ boson only leads to small quantitative changes in the results.
- [26] J. W. Rasul and T. Li, J. Phys. C **21**, 5119 (1988).
- [27] C. Lannert, M.P.A. Fisher, and T. Senthil, Phys. Rev. B **64**, 014518 (2001).
- [28] Philip Phillips, Rev. Mod. Phys. **82**, 1719 (2010).
- [29] F. Ronning, et. al. Science **282**, 2067 (1998).

SUPPLEMENTARY MATERIAL

Following the discussions in the main text, the partition function for the Hubbard model can be written as a functional integral over the quantum fields of the spin-carrying fermions $f_{i\sigma}$ and the slave bosons e_i , d_i , and $p_{i\sigma}$ introduced by Kotliar and Ruckenstein (KR) [1],

$$Z = \int \mathcal{D}[f, f^\dagger] \mathcal{D}[e, e^\dagger] \mathcal{D}[p, p^\dagger] \mathcal{D}[d, d^\dagger] \mathcal{D}[\lambda, \lambda_\sigma] e^{-\int_0^\beta \mathcal{L} d\tau}, \quad (\text{S1})$$

$$\mathcal{L} = \sum_i \left(e_i^\dagger \frac{\partial}{\partial \tau} e_i + d_i^\dagger \frac{\partial}{\partial \tau} d_i \right) + \sum_{i,\sigma} \left(p_{i\sigma}^\dagger \frac{\partial}{\partial \tau} p_{i\sigma} + f_{i\sigma}^\dagger \frac{\partial}{\partial \tau} f_{i\sigma} \right) + \hat{H}, \quad (\text{S2})$$

where \hat{H} is given in Eq. (1),

$$\hat{H} = -t \sum_{\langle i,j \rangle, \sigma} (\hat{z}_{i\sigma}^\dagger \hat{z}_{j\sigma} f_{i\sigma}^\dagger f_{j\sigma} + h.c.) + U \sum_i d_i^\dagger d_i + i \sum_i \lambda_i \hat{Q}_i + i \sum_{i,\sigma} \lambda_{i\sigma} \hat{Q}_{i\sigma} - \mu \sum_{i\sigma} f_{i\sigma}^\dagger f_{i\sigma}. \quad (\text{S3})$$

The functional integral over the Lagrange multipliers λ_i and $\lambda_{i\sigma}$ enforces the following local constraints

$$\hat{Q}_i = e_i^\dagger e_i + \sum_\sigma p_{i\sigma}^\dagger p_{i\sigma} + d_i^\dagger d_i - 1 = 0, \quad (\text{S4})$$

$$\hat{Q}_{i\sigma} = f_{i\sigma}^\dagger f_{i\sigma} - p_{i\sigma}^\dagger p_{i\sigma} - d_i^\dagger d_i = 0, \quad \forall \sigma. \quad (\text{S5})$$

The electron operator is written as $c_{i\sigma} = \hat{z}_{i\sigma} f_{i\sigma}$ where $\hat{z}_{i\sigma} = \hat{L}_{i\sigma} (e_i^\dagger p_{i\sigma} + p_{i\sigma}^\dagger d_i) \hat{R}_{i\sigma}$ with the normalization factors $\hat{R}_{i\sigma} = (1 - e_i^\dagger e_i - p_{i\sigma}^\dagger p_{i\sigma})^{-1/2}$ and $\hat{L}_{i\sigma} = (1 - d_i^\dagger d_i - p_{i\sigma}^\dagger p_{i\sigma})^{-1/2}$. Because of the projection-operator property of the bosons, the normal ordering of the square roots in the hopping renormalization factor $\hat{z}_{i\sigma}^\dagger \hat{z}_{i\sigma}$ is not an issue, as they can be translated directly into c -number boson fields in defining the functional integral.

A. Conventional slave-boson mean-field theory on honeycomb lattice

In the saddle-point approximation of the action considered by KR [1], which is referred to as the conventional slave-boson mean-field theory (SBMFT) or the Gutzwiller approximation, all the boson fields in Eq. (S3) are condensed: $e_i = e_0$, $d_i = d_0$, and $p_{i\sigma} = p_{0\sigma}$. For the honeycomb lattice Hubbard model at half-filling, the SBMFT was first studied by Frésard and Doll [2]. For the purpose of comparison, we plot the solutions in Fig. S1, which should be contrasted to the new saddle point solution in the presence of doublon-holon (D-H) binding shown in Fig. 2 in the main text. In the PM phase, the doublon density d_0^2 decreases linearly from 1/4 at $U = 0$ and vanishes at $U_{\text{BR}} \simeq 12.6t$, corresponding to the Brinkman-Rice metal-insulator transition (Fig. S1a). When magnetism is allowed, the semimetal to AF insulator transition takes place at $U_m \simeq 3.1t$ with the onset of the staggered magnetization $m = p_{0\uparrow}^2 - p_{0\downarrow}^2$. The doublon/holon (D/H) condensate density $d_0^2 = e_0^2$ remains nonzero for any finite Hubbard U , so the AF ordered phase in the SBMFT is a Slater insulator with coherent QP excitations above the AF insulating gap. The latter is determined by the energy gap in the fermion sector Ξ_f and scales linearly with $U - U_m$ as shown in Fig. S1b.

B. Doublon-holon binding and intersite correlations of the slave bosons

In order to construct the new saddle-point solution that takes into account the effects of D-H binding, we need to consider the dynamics of the bosons and the their intersite correlations. To do so, we introduce the operators for D-H pairing $\hat{\Delta}_{ij} = d_i e_j$ and the D/H hopping $\hat{\chi}_{ij}^d = d_i^\dagger d_j$, $\hat{\chi}_{ij}^e = e_i^\dagger e_j$ on the nearest-neighbor bonds, as well as the H/D density $\hat{n}_i^d = d_i^\dagger d_i$, $\hat{n}_i^e = e_i^\dagger e_i$ on each site. The hopping renormalization factor in Eq. (S3) can thus be written as

$$\hat{z}_{i\sigma}^\dagger \hat{z}_{j\sigma} = \hat{g}_{ij}^\sigma \left[(\hat{\chi}_{ij}^e)^\dagger p_{i\sigma}^\dagger p_{j\sigma} + \hat{\chi}_{ij}^d p_{i\sigma} p_{j\sigma}^\dagger + \hat{\Delta}_{ji} p_{i\sigma}^\dagger p_{j\sigma}^\dagger + \hat{\Delta}_{ij}^\dagger p_{i\sigma} p_{j\sigma} \right], \quad (\text{S6})$$

where the prefactor $\hat{g}_{ij}^\sigma = \hat{R}_{i\sigma}^\dagger \hat{L}_{i\sigma}^\dagger \hat{L}_{j\sigma} \hat{R}_{j\sigma}$ can be decoupled according to

$$\hat{g}_{ij}^\sigma = \left(\hat{Y}_{ij}^\sigma \hat{Y}_{ji}^\sigma \right)^{-1/2}, \quad \hat{Y}_{ij}^\sigma = (1 - p_{i\sigma}^\dagger p_{i\sigma})(1 - p_{j\sigma}^\dagger p_{j\sigma}) - \hat{n}_i^e (1 - p_{j\sigma}^\dagger p_{j\sigma}) - \hat{n}_j^d (1 - p_{i\sigma}^\dagger p_{i\sigma}) + |\hat{\Delta}_{ji}|^2. \quad (\text{S7})$$

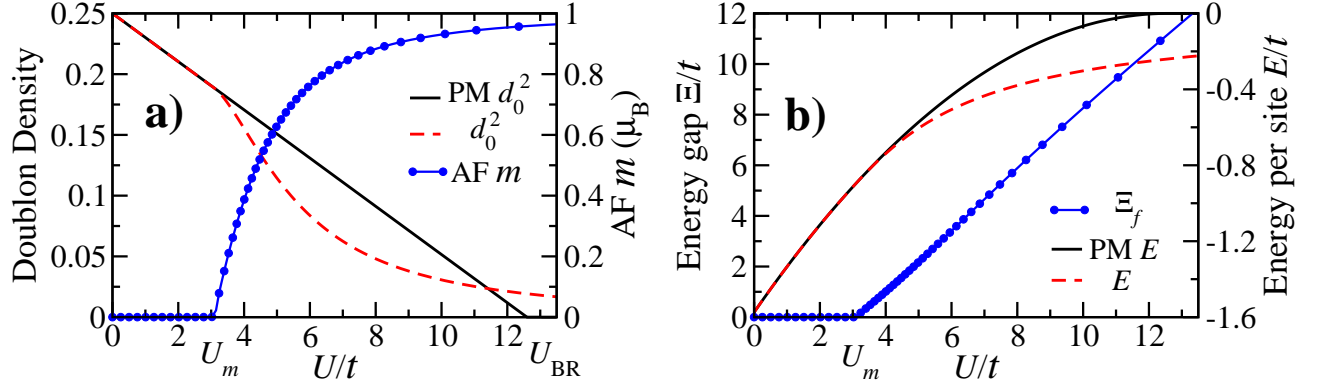


FIG. S1: (color online) Conventional SBMFT solution of the Hubbard model on half-filled honeycomb lattice. The Hubbard U dependence of (a) condensate density d_0^2 , staggered magnetization m , and (b) energy per site E and single-particle energy gap in the fermion sector Ξ_f .

Note that the above expression for \hat{g}_{ij}^σ involves explicitly the D-H pairing, but not the H/D hopping operators. As we shall show later, this choice of the decoupling ensures the correct description of the noninteracting limit at $U = 0$ where $\langle \hat{\Delta}_{ij} \rangle = d_0^2 = 1/4$ and $\langle \hat{\chi}_{ij}^e \rangle = e_0^2 = \langle \hat{\chi}_{ij}^d \rangle = d_0^2$ such that $z_{i\sigma} z_{j\sigma} = 1$.

The obvious challenge is how to build these correlations into the calculation of the partition function. Since they enter through the rather formidable factor $z_{i\sigma}^\dagger z_{j\sigma}$, the usual procedure of introducing the corresponding correlation fields $(\Delta_{ij}, \chi_{i,j}^e, \chi_{i,j}^d, n_i^d, n_i^e)$ as Hubbard-Stratonovich fields in the path-integral does not work here. We found that the difficulty can be overcome by introducing in the functional integral additional Lagrange multipliers in the corresponding channel, $\Delta_{ij}^v, \chi_{ij}^{d,v}, \chi_{ij}^{e,v}, \epsilon_i^{d,v}$, and $\epsilon_i^{e,v}$, such that the partition function becomes:

$$Z = \int \mathcal{D}[f, f^\dagger] \mathcal{D}[e, e^\dagger] \mathcal{D}[p, p^\dagger] \mathcal{D}[d, d^\dagger] \mathcal{D}[\lambda, \lambda_\sigma] \mathcal{D}[\Delta, \chi^d, \chi^e, n^d, n^e] \mathcal{D}[\Delta^v, \chi^{d,v}, \chi^{e,v}, \epsilon^{d,v}, \epsilon^{e,v}] e^{-\int_0^\beta \mathcal{L} d\tau}, \quad (\text{S8})$$

$$\mathcal{L} = \sum_i \left(e_i^\dagger \frac{\partial}{\partial \tau} e_i + d_i^\dagger \frac{\partial}{\partial \tau} d_i \right) + \sum_{i,\sigma} \left(p_{i\sigma}^\dagger \frac{\partial}{\partial \tau} p_{i\sigma} + f_{i\sigma}^\dagger \frac{\partial}{\partial \tau} f_{i\sigma} \right) + \hat{H}_{\text{DH}}, \quad (\text{S9})$$

where

$$\begin{aligned} \hat{H}_{\text{DH}} = & -t \sum_{\langle i,j \rangle, \sigma} (z_{i\sigma}^\dagger z_{j\sigma} f_{i\sigma}^\dagger f_{j\sigma} + h.c.) + U \sum_i d_i^\dagger d_i + i \sum_i \lambda_i \hat{Q}_i + i \sum_{i,\sigma} \lambda_{i\sigma} \hat{Q}_{i\sigma} - \mu \sum_{i\sigma} f_{i\sigma}^\dagger f_{i\sigma} \\ & - i \sum_{\langle i,j \rangle} \left[\chi_{ij}^{d,v} (d_i^\dagger d_j - \chi_{ij}^d) + \chi_{ij}^{e,v} (e_i^\dagger e_j - \chi_{ij}^e) + \Delta_{ij}^v (d_i e_j - \Delta_{ij}) + \Delta_{ji}^v (e_i d_j - \Delta_{ji}) + h.c. \right] \\ & + i \sum_i \left[\epsilon_i^{d,v} (d_i^\dagger d_i - n_i^d) + \epsilon_i^{e,v} (e_i^\dagger e_i - n_i^e) \right]. \end{aligned} \quad (\text{S10})$$

The effective H-D binding Hamiltonian in Eq. (S10) is just the one given in Eq. (2) of the main text, where the factor $z_{i\sigma} z_{j\sigma}$ has the operator form given in Eqs (S6) and (S7) but with the bosonic operators replaced by the corresponding correlation fields $(\Delta_{ij}, \chi_{i,j}^e, \chi_{i,j}^d, n_i^d, n_i^e)$. A few remarks are in order. (1) Eqs (S8)-(S10) provide an exact representation of the Hubbard model; carrying out formally the last two functional integrals in Eq. (S8) recovers the KR formulation given in Eqs (S1)-(S3). (2) In the above derivation, the intersite correlations of the $p_{i\sigma}$ bosons were not included explicitly for simplicity. The reason is that the $p_{i\sigma}$ bosons represent single-occupation; their density is large at half-filling. Thus, the $p_{i\sigma}$ bosons can be treated appropriately as condensed fields. We have verified that including their intersite correlations only leads to small quantitative changes in the results. (3) From the perspective of finding a saddle point solution of the action, the effective Hamiltonian H_{HD} in Eq. (S10) can be understood intuitively as a variational Hamiltonian describing the effects of intersite correlations of the doublons and holons, including that of H-D binding, as stated in the main text.

C. D-H binding saddle point solution and self-consistency equations

The saddle point solution corresponds to approximating the path integral in Eq. (S8) with a particular configuration of the quantum fields that minimizes the action. We consider here translation invariant paramagnetic (PM) and two-sublattice antiferromagnetic (AF) saddle point solutions on the honeycomb lattice at half-filling. The bond variables are real and isotropic and the symmetry requires $\Delta_{\langle ij \rangle} = \Delta_d$, $\chi_{\langle ij \rangle}^d = \chi_{\langle ij \rangle}^e = \chi_d$, $n_i^d = n_i^e = n_d$, and correspondingly, $i\Delta_{\langle ij \rangle}^v = \Delta_d^v$, $i\chi_{\langle ij \rangle}^{d,v} = i\chi_{\langle ij \rangle}^{e,v} = \chi_d^v$, $i\epsilon_i^{d,v} = i\epsilon_i^{e,v} = \epsilon_d^v$. Moreover $i\lambda_i = \lambda$, $i\lambda_{A\sigma} = i\lambda_{B\bar{\sigma}} = \lambda_\sigma$, and $p_{A0\sigma} = p_{B0\bar{\sigma}} = p_{0\sigma}$, where A and B denote the two sublattices on the honeycomb lattice. The normalization factor in Eq. (S7) takes on the following saddle point value on the nearest neighbor bonds,

$$g_{\langle ij \rangle}^\sigma = g = \prod_\sigma Y_\sigma^{-1/2}, \quad \text{with } Y_\sigma = 1 - 2n_d - 2p_{0\sigma}^2 + 2p_{0\sigma}^2 n_d + p_{0\sigma}^4 + \Delta_d^2. \quad (\text{S11})$$

Substituting these quantities into Eq. (S10), we obtain the saddle point Hamiltonian in the main text,

$$\hat{H}_{\text{HD}}^{\text{sp}} = \hat{H}_f + \hat{H}_d + 12N(\chi_d^v \chi_d + \Delta_d^v \Delta_d) - 2N(\epsilon_d^v + 2\lambda) - 4N\epsilon_d^v n_d + 2N \sum_\sigma (\lambda - \lambda_\sigma) p_{0\sigma}^2 \quad (\text{S12})$$

on the honeycomb lattice with $2N$ number of sites, where \hat{H}_f and \hat{H}_d describe the energy spectrum in the fermion and boson sectors. In momentum space, the fermion part is given by

$$\hat{H}_f = \sum_{\mathbf{k}, \sigma} \begin{bmatrix} f_{A\mathbf{k}\sigma} \\ f_{B\mathbf{k}\sigma} \end{bmatrix}^\dagger \begin{bmatrix} \lambda_\sigma - \mu & -\chi_f^v \eta_{\mathbf{k}} \\ -\chi_f^v \eta_{\mathbf{k}}^* & \lambda_{\bar{\sigma}} - \mu \end{bmatrix} \begin{bmatrix} f_{A\mathbf{k}\sigma} \\ f_{B\mathbf{k}\sigma} \end{bmatrix}, \quad (\text{S13})$$

where the A-B sublattice hopping amplitude $\chi_f^v = tg \left[2p_{0\uparrow} p_{0\downarrow} \chi_d + (p_{0\uparrow}^2 + p_{0\downarrow}^2) \Delta_d \right]$ and $\eta_{\mathbf{k}} = \exp(ik_y) + 2\cos(\sqrt{3}k_x/2) \exp(-ik_y/2)$ on the honeycomb lattice. At half-filling, the particle-hole symmetry requires

$$\mu = \frac{U}{2}, \quad \text{and } \lambda_\uparrow + \lambda_\downarrow = U. \quad (\text{S14})$$

We can thus write $\lambda_\sigma = \mu - \sigma\varepsilon$, where $\varepsilon = (\lambda_\uparrow - \lambda_\downarrow)/2$ is equal to zero in the PM phase, but nonzero in the AF phase. The fermion action can therefore be diagonalized by the transformation

$$\begin{bmatrix} f_{A\mathbf{k}\sigma}(\tau) \\ f_{B\mathbf{k}\sigma}(\tau) \end{bmatrix} = \begin{bmatrix} \tilde{u}_{A\mathbf{k}\sigma} & \tilde{v}_{A\mathbf{k}\sigma} \\ \tilde{u}_{B\mathbf{k}\sigma} & \tilde{v}_{B\mathbf{k}\sigma} \end{bmatrix} \begin{bmatrix} \gamma_{\sigma-}(\mathbf{k}) e^{-\tau E_-^f(\mathbf{k})} \\ \gamma_{\sigma+}(\mathbf{k}) e^{-\tau E_+^f(\mathbf{k})} \end{bmatrix}, \quad (\text{S15})$$

where $\gamma_{\sigma\pm}(\mathbf{k})$ are the fermion quasiparticles (QPs) with eigenstate energy dispersions:

$$E_\pm^f(\mathbf{k}) = \pm E_{\mathbf{k}}^f, \quad E_{\mathbf{k}}^f = \sqrt{\varepsilon^2 + |\chi_f^v \eta_{\mathbf{k}}|^2}. \quad (\text{S16})$$

The corresponding eigenvectors, satisfying $\tilde{u}_{A\mathbf{k}\sigma} = \tilde{v}_{B\mathbf{k}\sigma}^* = \tilde{u}_{\mathbf{k}\sigma}$ and $\tilde{v}_{A\mathbf{k}\sigma} = -\tilde{u}_{B\mathbf{k}\sigma}^* = \tilde{v}_{\mathbf{k}\sigma}$, are determined by

$$|\tilde{u}_{\mathbf{k}\sigma}|^2 = \frac{1}{2} \left(1 + \frac{\sigma\varepsilon}{E_{\mathbf{k}}^f} \right), \quad |\tilde{v}_{\mathbf{k}\sigma}|^2 = \frac{1}{2} \left(1 - \frac{\sigma\varepsilon}{E_{\mathbf{k}}^f} \right), \quad \tilde{u}_{\mathbf{k}\sigma} \tilde{v}_{\mathbf{k}\sigma} = -\frac{\chi_f^v \eta_{\mathbf{k}}}{2E_{\mathbf{k}}^f}. \quad (\text{S17})$$

The fermion spectrum in Eq. (S16) shows that in the semimetallic PM phase ($\varepsilon = 0$), the QP has Dirac dispersion with a renormalized velocity χ_f^v ; whereas a fermion energy gap

$$\Xi_f = 2|\varepsilon| \quad (\text{S18})$$

opens at the Dirac points in the AF ordered phase. The fermion hopping and density per spin are readily obtained,

$$\chi_f = \chi_{f\sigma} = \frac{1}{6N} \sum_{\mathbf{k}} \langle \eta_{\mathbf{k}} f_{A\mathbf{k}\sigma}^\dagger f_{B\mathbf{k}\sigma} + h.c. \rangle = \frac{1}{6N} \chi_f^v \sum_{\mathbf{k}} \frac{|\eta_{\mathbf{k}}|^2}{E_{\mathbf{k}}^f}, \quad (\text{S19})$$

$$n_\sigma^f = \frac{1}{N} \sum_{\mathbf{k}} \langle f_{A\mathbf{k}\sigma}^\dagger f_{A\mathbf{k}\sigma} \rangle = \frac{1}{N} \sum_{\mathbf{k}} \langle f_{B\mathbf{k}\bar{\sigma}}^\dagger f_{B\mathbf{k}\bar{\sigma}} \rangle = \frac{1}{2N} \sum_{\mathbf{k}} \left(1 + \frac{\sigma\varepsilon}{E_{\mathbf{k}}^f} \right). \quad (\text{S20})$$

The spectrum of the charged (d - e) boson sector is determined by \hat{H}_d in the saddle point Hamiltonian (S12), which can be written in a 4×4 matrix form in momentum space,

$$\hat{H}_d = \sum_{\mathbf{k}} \begin{bmatrix} d_{A\mathbf{k}} \\ d_{B\mathbf{k}} \\ e_{B\bar{\mathbf{k}}}^\dagger \\ e_{A\bar{\mathbf{k}}}^\dagger \end{bmatrix}^\dagger \begin{bmatrix} \epsilon_d^\vee + \lambda & -\chi_d^\vee \eta_{\mathbf{k}} & -\Delta_d^\vee \eta_{\mathbf{k}} & 0 \\ -\chi_d^\vee \eta_{\mathbf{k}}^* & \epsilon_d^\vee + \lambda & 0 & -\Delta_d^\vee \eta_{\mathbf{k}}^* \\ -\Delta_d^\vee \eta_{\mathbf{k}}^* & 0 & \epsilon_d^\vee + \lambda & -\chi_d^\vee \eta_{\mathbf{k}}^* \\ 0 & -\Delta_d^\vee \eta_{\mathbf{k}} & -\chi_d^\vee \eta_{\mathbf{k}} & \epsilon_d^\vee + \lambda \end{bmatrix} \begin{bmatrix} d_{A\mathbf{k}} \\ d_{B\mathbf{k}} \\ e_{B\bar{\mathbf{k}}}^\dagger \\ e_{A\bar{\mathbf{k}}}^\dagger \end{bmatrix}, \quad (\text{S21})$$

where the relations due to particle-hole symmetry in Eq. (S14) have been applied. The bosonic action can be diagonalized by the following Bogoliubov transformation,

$$\begin{bmatrix} d_{A\mathbf{k}}(\tau) \\ d_{B\mathbf{k}}(\tau) \\ e_{B\bar{\mathbf{k}}}^\dagger(\tau) \\ e_{A\bar{\mathbf{k}}}^\dagger(\tau) \end{bmatrix} = \begin{bmatrix} u_{A\mathbf{k}+} & u_{A\mathbf{k}-} & v_{A\mathbf{k}+} & v_{A\mathbf{k}-} \\ u_{B\mathbf{k}+} & u_{B\mathbf{k}-} & v_{B\mathbf{k}+} & v_{B\mathbf{k}-} \\ v_{B\mathbf{k}+} & v_{B\mathbf{k}-} & u_{B\mathbf{k}+} & u_{B\mathbf{k}-} \\ v_{A\mathbf{k}+} & v_{A\mathbf{k}-} & u_{A\mathbf{k}+} & u_{A\mathbf{k}-} \end{bmatrix} \begin{bmatrix} a_+(\mathbf{k})e^{-\tau E_+^d(\mathbf{k})} \\ a_-(\mathbf{k})e^{-\tau E_-^d(\mathbf{k})} \\ b_+^\dagger(\mathbf{k})e^{\tau E_+^d(\mathbf{k})} \\ b_-^\dagger(\mathbf{k})e^{\tau E_-^d(\mathbf{k})} \end{bmatrix}, \quad (\text{S22})$$

where $a_\pm(\mathbf{k})$ and $b_\pm(\mathbf{k})$ are four bosonic QPs with two branches of doubly degenerate eigenvalues

$$E_\pm^d(\mathbf{k}) = \sqrt{(\epsilon_d^\vee + \lambda \pm |\chi_d^\vee \eta_{\mathbf{k}}|)^2 - |\Delta_d^\vee \eta_{\mathbf{k}}|^2}, \quad (\text{S23})$$

under the condition $\epsilon_d^\vee + \lambda \geq 3(|\chi_d^\vee| + |\Delta_d^\vee|)$ that ensures the eigenvalues to be real and positive for bosons. The factor of 3 in the latter arises from the maximum value of $\eta_{\mathbf{k}}$. The energy gap in the d - e boson sector is thus given by

$$\Xi_d = 2\sqrt{(\epsilon_d^\vee + \lambda - 3|\chi_d^\vee|)^2 - (3|\Delta_d^\vee|)^2}. \quad (\text{S24})$$

Analytical expressions of the corresponding eigenvectors $\{u, v\}$ are more lengthy than informative and thus omitted here. In practice, they are computed numerically.

The saddle point solution for a given Hubbard U can be determined by solving the set of self-consistency equations derived from minimizing the energy with respect to the ten variables $\{\chi_d, \Delta_d, n_d, p_{0\sigma}, \varepsilon, \lambda, \chi_d^\vee, \Delta_d^\vee, \epsilon_d^\vee\}$:

$$\langle \partial \hat{H}_{\text{MF}} / \partial \chi_d \rangle = 0 \mapsto \chi_d^\vee = 2tg p_{0\uparrow} p_{0\downarrow} \chi_f, \quad (\text{S25})$$

$$\langle \partial \hat{H}_{\text{MF}} / \partial \Delta_d \rangle = 0 \mapsto \Delta_d^\vee = g\chi_f \sum_{\sigma} (tp_{0\sigma}^2 - g\Delta_d \chi_f^\vee Y_{\sigma}), \quad (\text{S26})$$

$$\langle \partial \hat{H}_{\text{MF}} / \partial n_d \rangle = 0 \mapsto \epsilon_d^\vee = -3g^2 \chi_f \chi_f^\vee \sum_{\sigma} (1 - p_{0\bar{\sigma}}^2) Y_{\sigma}, \quad (\text{S27})$$

$$\langle \partial \hat{H}_{\text{eff}} / \partial p_{0\sigma} \rangle = 0 \mapsto [\lambda - U/2 + \sigma\epsilon - 6g^2 \chi_f \chi_f^\vee Y_{\bar{\sigma}} (1 - n_d - p_{0\sigma}^2)] p_{0\sigma} = 6tg \chi_f (\chi_d p_{0\bar{\sigma}} + \Delta_d p_{0\sigma}), \quad (\text{S28})$$

$$\langle \partial \hat{H}_{\text{eff}} / \partial \varepsilon \rangle = 0 \mapsto m = p_{0\uparrow}^2 - p_{0\downarrow}^2 = n_{\uparrow}^f - n_{\downarrow}^f, \quad (\text{S29})$$

$$\langle \partial \hat{H}_{\text{eff}} / \partial \lambda \rangle = 0 \mapsto 2n_d + p_{0\uparrow}^2 + p_{0\downarrow}^2 = 1, \quad (\text{S30})$$

$$\langle \partial \hat{H}_{\text{MF}} / \partial \chi_d^\vee \rangle = 0 \mapsto \chi_d = d_0^2 + \frac{1}{6N} \sum_{\mathbf{k}}' \langle \eta_{\mathbf{k}} d_{A\mathbf{k}}^\dagger d_{B\mathbf{k}} + h.c. \rangle, \quad (\text{S31})$$

$$\langle \partial \hat{H}_{\text{MF}} / \partial \Delta_d^\vee \rangle = 0 \mapsto \Delta_d = d_0^2 + \frac{1}{6N} \sum_{\mathbf{k}}' \langle \eta_{\mathbf{k}}^* d_{A\mathbf{k}} e_{B\bar{\mathbf{k}}} + h.c. \rangle, \quad (\text{S32})$$

$$\langle \partial \hat{H}_{\text{MF}} / \partial \epsilon_d^\vee \rangle = 0 \mapsto n_d = d_0^2 + \frac{1}{2N} \sum_{\alpha=\{A,B\}} \sum_{\mathbf{k}}' \langle d_{\alpha\mathbf{k}} d_{\alpha\mathbf{k}} \rangle. \quad (\text{S33})$$

It is instructive to examine the last three equations for the D/H hopping, D-H pairing, and the D/H density fields. The closing of the gap Ξ_d in the boson spectrum gives rise to a zero energy mode at the $\Gamma = (0,0)$ point [see Fig. (S3)] whose occupation enables the single-boson condensates $d_0 = e_0 \neq 0$. Thus, whenever $d_0 \neq 0$, the Γ point is removed from the momentum summations in Eqs (S31, S32, S33) such that the remaining summation over k comes solely from the contributions introduced by the D-H binding. Accordingly, the solutions to this set of self-consistency equations (S25-S33) must be searched with two attempts: (i) assume $d_0 = 0$ and solve these ten equations for the ten unknown

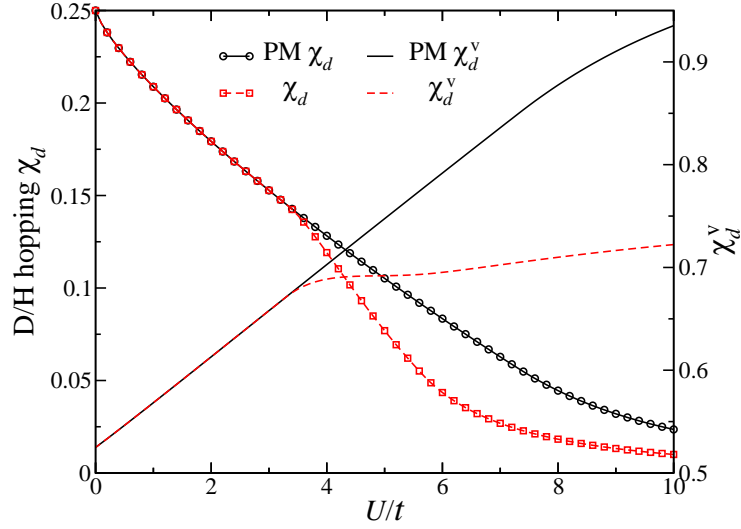


FIG. S2: (color online) The Hubbard U dependence of the variational parameter χ_d^v and the D/H hopping parameter χ_d . The corresponding results in the restricted PM phase are also plotted for comparison.

variables and (ii) assume a nonzero d_0 . The latter introduces one more unknown variable (d_0), while the required condition for the existence of the zero energy mode gives us the extra equation

$$\epsilon_d^v + \lambda = 3(|\chi_d^v| + |\Delta_d^v|), \quad (\text{S34})$$

and thus one can determine the state by solving the eleven self-consistency equations for the eleven unknown variables. If several solutions are found for a given U , the one with the lowest energy should be chosen as the ground state. In practice, we found only one solution for any given U .

Since most of the results have been discussed in the main text, we provide here only a few additional details. In Fig. S2, the self-consistently determined variational parameter χ_d^v and the D/H hopping parameter χ_d are plotted as a function of U in the ground state, as well as in the restricted PM phase. At $U = 0$, $\chi_d^v \neq 0$. However, since $\Delta_d^v = 0$ (Fig. 2a in the main text) at $U = 0$ leads to $\chi_d = 1/4$, the noninteracting limit is correctly recovered. The variational parameter χ_d^v increases and the D/H hopping field χ_d decreases monotonically with increasing U . Note that χ_d remains finite for any finite U and evolves analytically across the Mott transitions at $U_c = 7.3t$ in the PM phase and at the AF \rightarrow AF* transition at $U_{c2} = 5.7t$ in the ground state, as does the D-H pairing field Δ_d shown in Fig. 2a in the main text. As discussed in the main text, the nonzero values of χ_d and Δ_d ensure that both the uniform and the staggered components of the charge U(1) symmetry are broken and the associated gauge fields are massive in the electron-fractionalized AF* phase.

It is also enlightening to discuss the boson energy spectrum and the integrated spectral function (ISF) in the doublon/holon sector, which are plotted in Fig. S3 in the PM phase at several values of U across the Mott transition. Their behaviors in the AF and AF* phases are similar. At a given U , the boson spectrum shows the two doubly-degenerate dispersive branches obtained in Eq. (S23) and Eq. (S24). There are several noteworthy features. (i) Both dispersive branches are flat around the M point of the hexagonal Brillouin zone, leading to the two Van Hove singularities in the boson ISF shown in the right panel of Fig. S3. (ii) The two branches cross and produce the Dirac-cone-like behavior at the K point at a finite energy that increases with increasing U , leading to the V-shaped density of states. Remarkably, (i) and (ii) combine to form a Dirac-cone like dispersion that is similar and can be regarded as a “ghost” band of the bare electron dispersion carried by the excitations of the doublon/holon complex. This property was pointed out in the systematic large- N expansion study of the t - J model for doped Mott insulators [4]. It is remarkable that the ghost Dirac-cone feature manifests itself in the broad peak-dip-peak structure in the ISF of the physical electrons shown in Fig. 3 of the main text, which can be identified as the holon-doublon excitations. (iii) The low energy properties of the boson dispersion near the Γ -point is intriguing and important. For $U < U_c$, i.e. on the metallic side of the Mott transition, the lower energy branch is gapless, i.e. $E_-^d(\mathbf{k}) = 0$ and disperses linearly away from the Γ point. The existence of the zero-energy mode, together with the vanishing of the DOS/ISF $N(E)$, enables the finite-temperature D/H condensation in a two-dimensional system such that $d_0 \neq 0$ at zero temperature. In contrast, for $U > U_c$, or $U > U_{c2}$ when magnetism is allowed, a charge gap $\Xi_d \neq 0$ opens up at the Γ point,

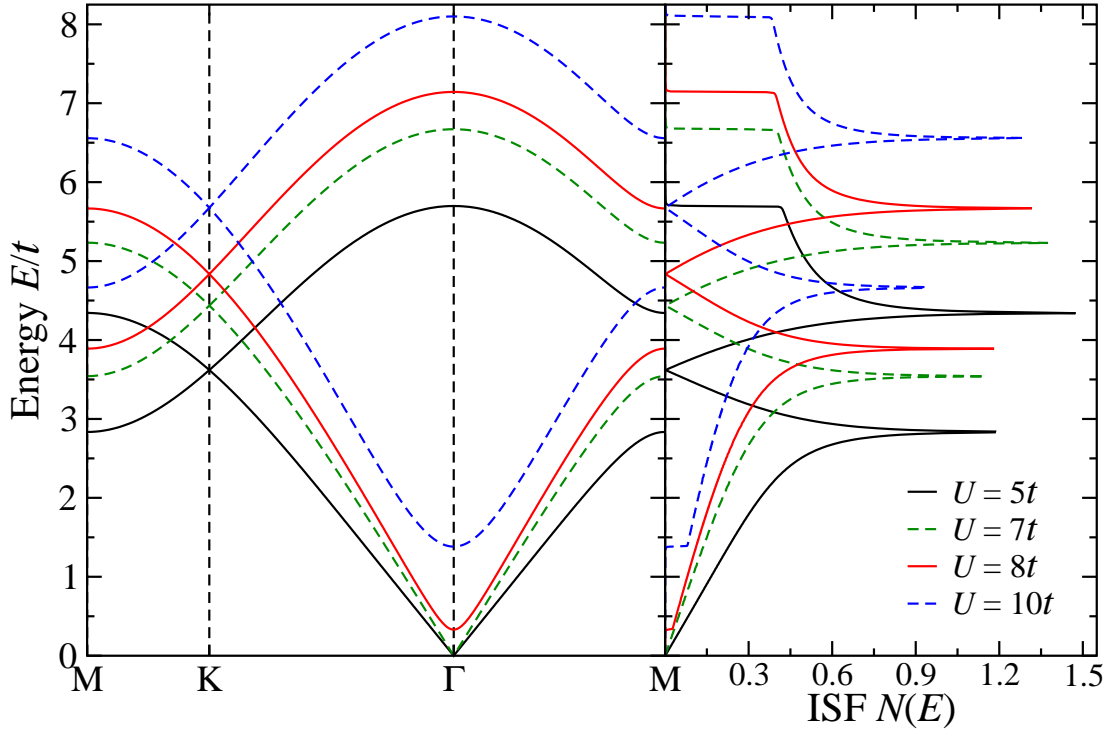


FIG. S3: (color online) The energy spectrum along high symmetry directions (left panel) and the ISF (right panel) of the doublons and holons in the PM phases at several Hubbard U across the Mott transition at $U_c = 7.3t$. The contribution from the condensate on the semimetallic side of the Mott transition is not shown.

indicating the emergence of the Mott insulating or the AF* phase with complete suppression of the D/H condensate. Note also that the gapped $E_-^d(\mathbf{k})$ is parabolic near Γ , giving rise to a finite ISF $N(E)$ at the band bottom.

D. Integrated spectral function for the physical electrons

We provide here additional discussions on the calculation of the integrated spectral function (ISF), which is also the tunneling density of states (DOS), for the physical electrons. The electron spectral function is defined in terms of the retarded single-particle Green's function $G_\sigma(\mathbf{k}, \omega)$ [3]:

$$A_\alpha(\mathbf{k}, \omega) = -\text{Im} \sum_\sigma G_{\alpha\sigma}(\mathbf{k}, \omega), \quad G_{\alpha\sigma}(\mathbf{k}, \tau) = -\langle T_\tau c_{\alpha\mathbf{k}\sigma}(\tau) c_{\alpha\mathbf{k}\sigma}^\dagger(0) \rangle. \quad (\text{S35})$$

Because of the presence of two sublattices on the honeycomb lattice, the Green's function and the spectral function acquire an additional sublattice index α in the above equations. The ISF or the tunneling DOS, is given by

$$N_\alpha(\omega) = -\sum_{\mathbf{k}, \sigma} \text{Im} G_{\alpha\sigma}(\mathbf{k}, \omega) = -\sum_{\mathbf{k}, \sigma} \text{Im} \int_0^\beta d\tau e^{i\omega\tau} G_{\alpha\sigma}(\mathbf{k}, \tau). \quad (\text{S36})$$

As discussed in the main text, the electron operator is composed of $c_{i\sigma} = \hat{L}_{i\sigma}(e_i^\dagger p_{i\sigma} + p_{i\bar{\sigma}}^\dagger d_i) \hat{R}_{i\sigma} f_{i\sigma}$ in the KR slave boson formulation [1]. Within our saddle point solution, we keep the product of the fermion and slave-boson operators and approximate the normalization factors by the saddle point average. The electron operator in momentum space is given by,

$$c_{\alpha\mathbf{k}\sigma} = r_{\alpha\sigma} \sum_{\mathbf{q}, \mathbf{q}'} (e_{\alpha\bar{\mathbf{q}}}^\dagger p_{\alpha\mathbf{q}'\sigma} + p_{\alpha\bar{\mathbf{q}}'\sigma}^\dagger d_{\mathbf{q}}) f_{\alpha, \mathbf{k}-\mathbf{q}-\mathbf{q}', \sigma}, \quad (\text{S37})$$

with the normalization factor

$$r_{\alpha\sigma} = \langle \hat{L}_{\alpha\sigma} \hat{R}_{\alpha\sigma} \rangle = (1 - n_d - p_{\alpha 0\sigma}^2)^{-1/2} (1 - n_d - p_{\alpha 0\bar{\sigma}}^2)^{-1/2}. \quad (\text{S38})$$

The electron Green's function in Eq. (S35) is therefore given by

$$G_{\alpha\sigma}(\mathbf{k}, \tau) = r_{\alpha\sigma}^2 \sum_{\mathbf{q}, \mathbf{q}'} \Lambda_{\alpha\sigma}(\mathbf{q}, \mathbf{q}', \tau) G_{\alpha\sigma}^f(\mathbf{k} - \mathbf{q} - \mathbf{q}', \tau), \quad (\text{S39})$$

where G^f is the f_σ -fermion Green's function and Λ involves the normal and anomalous (due to pairing) Green's functions of the bosons

$$\begin{aligned} \Lambda_{\alpha\sigma}(\mathbf{q}, \mathbf{q}', \tau) = & \langle T_\tau e_{\alpha\bar{\mathbf{q}}}^\dagger(\tau) e_{\alpha\bar{\mathbf{q}}}(0) \rangle \langle T_\tau p_{\alpha\mathbf{q}'\sigma}(\tau) p_{\alpha\mathbf{q}'\sigma}^\dagger(0) \rangle + \langle T_\tau d_{\alpha\mathbf{q}}(\tau) d_{\alpha\mathbf{q}}^\dagger(0) \rangle \langle T_\tau p_{\alpha\bar{\mathbf{q}}'\bar{\sigma}}^\dagger(\tau) p_{\alpha\bar{\mathbf{q}}'\bar{\sigma}}(0) \rangle \\ & + \langle T_\tau e_{\alpha\bar{\mathbf{q}}}^\dagger(\tau) d_{\alpha\bar{\mathbf{q}}}^\dagger(0) \rangle \langle T_\tau p_{\alpha\mathbf{q}'\sigma}(\tau) p_{\alpha\bar{\mathbf{q}}'\bar{\sigma}}(0) \rangle + \langle T_\tau d_{\alpha\mathbf{q}}(\tau) e_{\alpha\bar{\mathbf{q}}}(0) \rangle \langle T_\tau p_{\alpha\bar{\mathbf{q}}'\bar{\sigma}}^\dagger(\tau) p_{\alpha\mathbf{q}'\sigma}^\dagger(0) \rangle. \end{aligned} \quad (\text{S40})$$

The ISF of the physical electrons in Eq. (S36) becomes

$$N_\alpha(\omega) = - \sum_{\mathbf{k}, \sigma} r_{\alpha\sigma}^2 \text{Im} \int_0^\beta d\tau e^{i\omega\tau} \Lambda_{\alpha\sigma}(\tau) G_{\alpha\sigma}^f(\mathbf{k}, \tau), \quad \text{where} \quad \Lambda_{\alpha\sigma}(\tau) = \sum_{\mathbf{q}, \mathbf{q}'} \Lambda_{\alpha\sigma}(\mathbf{q}, \mathbf{q}', \tau). \quad (\text{S41})$$

In frequency space, the spectral function thus involves convolutions of the boson Green's functions with that of the f_σ -fermions, and as a result, the single-particle energy gap in the DOS of the physical electron is the sum of those in the fermion and boson sectors: $\Xi = \Xi_f + \Xi_d + \Xi_p$. Within our theory, the p_σ -boson condensate exists over the entire phase diagram and thus $\Xi_p = 0$.

It is instructive to write each boson operator as the sum of the condensate and fluctuations: $b_{\mathbf{k}}^{(\dagger)} = b_0 \delta_{\mathbf{k}} + \tilde{b}_{\mathbf{k}}^{(\dagger)}$, where b stands for the (d, e, p_σ) bosons. Although this is not necessary, doing so facilitates well the following discussions of the coherent and incoherent contributions to the electron spectral function. Note that the fluctuations $\tilde{b}_{\mathbf{k}}^{(\dagger)}$ are boson operators, obeying boson commutation relations and the energy spectrum discussed above. Thus, the normal and anomalous boson Green's functions can be written as

$$\langle T_\tau b_{\mathbf{k}}^{(\dagger)}(\tau) b_{\mathbf{q}}^{(\dagger)}(0) \rangle = b_0 b_0' \delta_{\mathbf{k}\mathbf{q}} + \langle T_\tau \tilde{b}_{\mathbf{k}}^{(\dagger)}(\tau) \tilde{b}_{\mathbf{q}}^{(\dagger)}(0) \rangle. \quad (\text{S42})$$

Decomposing the condensate and fluctuation contributions in Eq. (S40) this way and keeping the leading order fluctuations involving a single boson Green's function, the $\Lambda_{\alpha\sigma}(\tau)$ in Eq. (S41) becomes

$$\Lambda_{\alpha\sigma}(\tau) = \Lambda_{\alpha\sigma}^{\text{cond}}(\tau) + \Lambda_{\alpha\sigma}^{\text{fluc}}(\tau),$$

where the condensate part $\Lambda_{\alpha\sigma}^{\text{cond}}(\tau) = d_0^2 (p_{0\uparrow} + p_{0\downarrow})^2$, and the fluctuation part

$$\begin{aligned} \Lambda_{\alpha\sigma}^{\text{fluc}}(\tau) = & d_0^2 \sum_{\mathbf{q}} \left[\langle T_\tau \tilde{p}_{\alpha\mathbf{q}\sigma}(\tau) \tilde{p}_{\alpha\mathbf{q}\sigma}^\dagger(0) \rangle + \langle T_\tau \tilde{p}_{\alpha\bar{\mathbf{q}}\bar{\sigma}}^\dagger(\tau) \tilde{p}_{\alpha\bar{\mathbf{q}}\bar{\sigma}}(0) \rangle \right] + p_{\alpha 0\sigma}^2 \sum_{\mathbf{q}} \langle T_\tau \tilde{e}_{\alpha\bar{\mathbf{q}}}^\dagger(\tau) \tilde{e}_{\alpha\bar{\mathbf{q}}}(0) \rangle + p_{\alpha\bar{\sigma}}^2 \sum_{\mathbf{q}} \langle T_\tau \tilde{d}_{\alpha\mathbf{q}}(\tau) \tilde{d}_{\alpha\mathbf{q}}^\dagger(0) \rangle \\ & + p_{0\uparrow} p_{0\downarrow} \sum_{\mathbf{q}} \left[\langle T_\tau \tilde{e}_{\alpha\bar{\mathbf{q}}}^\dagger(\tau) \tilde{d}_{\alpha\mathbf{q}}^\dagger(0) \rangle + \langle T_\tau \tilde{d}_{\alpha\mathbf{q}}(\tau) \tilde{e}_{\alpha\bar{\mathbf{q}}}(0) \rangle \right]. \end{aligned} \quad (\text{S43})$$

Correspondingly, the ISF in Eq. (S41) can be written as

$$N_\alpha(\omega) = N_\alpha^{\text{coh}}(\omega) + N_\alpha^{\text{incoh}}(\omega), \quad (\text{S44})$$

with

$$N_\alpha^{\text{coh(incoh)}}(\omega) = - \sum_{\mathbf{k}, \sigma} r_{\alpha\sigma}^2 \text{Im} \int_0^\beta d\tau e^{i\omega\tau} \Lambda_{\alpha\sigma}^{\text{cond(fluc)}}(\tau) G_{\alpha\sigma}^f(\mathbf{k}, \tau). \quad (\text{S45})$$

The coherent part of the ISF comes from the single-boson condensates that recombine the charge and spin degrees of freedom, leading to coherent quasiparticle (QP) excitations associated with the coherence peaks in the ISF.

The f_σ -fermion Green's function is easy to compute in terms of the fermion QP wave functions in Eq. (S15),

$$G_{\alpha\sigma}^f(\mathbf{k}, \tau) = - \langle T_\tau f_{\alpha\mathbf{k}\sigma}(\tau) f_{\alpha\mathbf{k}\sigma}^\dagger(0) \rangle = - |\tilde{u}_{\alpha\mathbf{k}\sigma}|^2 n_F(E_{\mathbf{k}}^f) e^{\tau E_{\mathbf{k}}^f} - |\tilde{v}_{\alpha\mathbf{k}\sigma}|^2 n_F(-E_{\mathbf{k}}^f) e^{-\tau E_{\mathbf{k}}^f}, \quad (\text{S46})$$

where $n_F(E) = 1/[\exp(\beta E) + 1]$ is the Fermi distribution function and $\beta = 1/k_B T$ with T the temperature. As a result, the coherent contribution to the ISF is given by

$$N_\alpha^{\text{coh}}(\omega) = \pi d_0^2 (p_{0\uparrow} + p_{0\downarrow})^2 \sum_{\mathbf{k}, \sigma} r_{\alpha\sigma}^2 \left[|\tilde{u}_{\alpha\mathbf{k}\sigma}|^2 \delta(\omega + E_{\mathbf{k}}^f) + |\tilde{v}_{\alpha\mathbf{k}\sigma}|^2 \delta(\omega - E_{\mathbf{k}}^f) \right]. \quad (\text{S47})$$

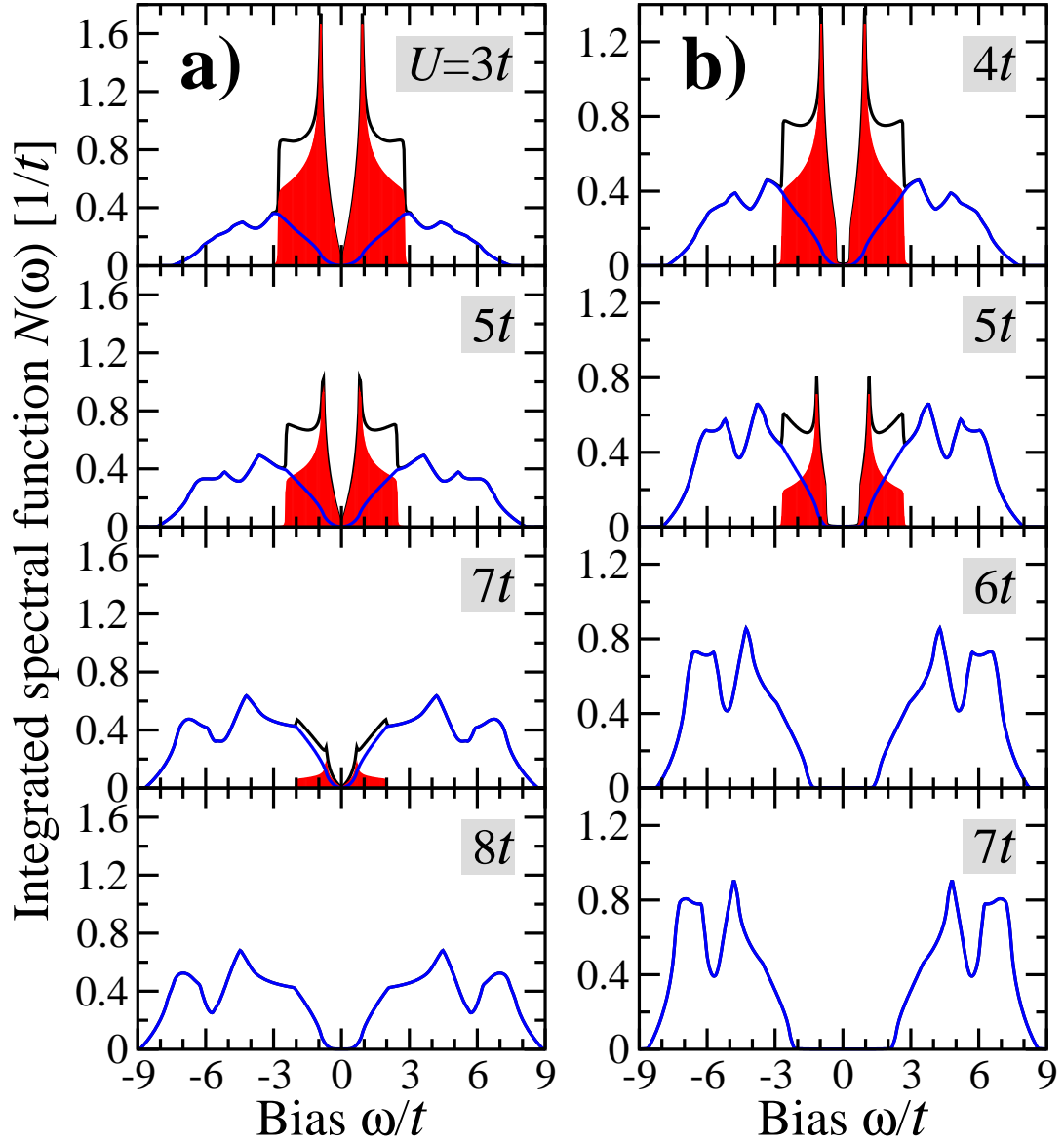


FIG. S4: (color online) The coherent (shaded red areas), incoherent (blue solid lines), and the total (black solid lines) integrated spectral function at different values of Hubbard U in (a) the restricted PM phase where the Mott transition is at $U_c \simeq 7.3t$ and (b) the ground state where the AF to AF* transition is at $U_{c2} \simeq 5.7t$.

Because $|\tilde{u}_{\alpha\mathbf{k}\sigma}|^2 + |\tilde{v}_{\alpha\mathbf{k}\sigma}|^2 = 1$, the total coherent spectral weight of the ISF is

$$W_{\alpha}^{\text{coh}} = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} N_{\alpha}^{\text{coh}}(\omega) = \frac{1}{2} d_0^2 (p_{0\uparrow} + p_{0\downarrow})^2 \sum_{\sigma} r_{\alpha\sigma}^2.$$

In the noninteracting limit at $U = 0$, the saddle point solution has $d_0 = p_{0\sigma} = 1/2$ and $r_{\alpha\sigma} = 2$, such that the total coherent weight $W_{\alpha}^{\text{coh}} = 1$ with the complete suppression of the incoherent part of the spectral function. In the opposite limit on the Mott insulator side, where the doublon/holon condensate $d_0 = e_0 = 0$, the coherent spectral function is completely suppressed $W_{\alpha}^{\text{coh}} = 0$ and the entire spectral weight must come from the incoherent part of the spectral function.

In the incoherent part of the ISF, defined in Eqs (S45) and (S43), the convolution of the boson and fermion Green's functions gives broad spectral features. Since in our saddle point solution, the p_{σ} bosons are fully condensed and their fluctuations were ignored for simplicity, the question arises as to how to evaluate the corresponding Green's functions in Eq. (S43). Naively, one would simply ignore these fluctuations; but we found that doing so would produce

an unphysical result of a nonzero incoherent contribution to the ISF in the noninteracting limit. The problem can be resolved by realizing that at the saddle point level, the local constraint in Eq. (S4) is only satisfied on average, i.e. $\langle \hat{Q}_{i\alpha} \rangle = 0$. When fluctuations are considered, a consistent condition imposed by the constraint is $\langle T_\tau \hat{Q}_\alpha(\tau) \hat{Q}_\alpha(0) \rangle = 0$ where $\hat{Q}_\alpha = (1/N) \sum_{i \in \alpha} \hat{Q}_{i\alpha}$ with N the number of α -sublattice sites. Evaluating the latter leads to the following expression

$$d_0^2 \sum_{\mathbf{k}} \left[\langle T_\tau \tilde{e}_{\alpha\bar{\mathbf{k}}}^\dagger(\tau) \tilde{e}_{\alpha\bar{\mathbf{k}}}(0) \rangle + \langle T_\tau \tilde{d}_{\alpha\mathbf{k}}(\tau) \tilde{d}_{\alpha\mathbf{k}}^\dagger(0) \rangle + \langle T_\tau \tilde{e}_{\alpha\bar{\mathbf{k}}}^\dagger(\tau) \tilde{d}_{\alpha\mathbf{k}}^\dagger(0) \rangle + \langle T_\tau \tilde{d}_{\alpha\mathbf{k}}(\tau) \tilde{e}_{\alpha\bar{\mathbf{k}}}(0) \rangle \right] \\ + \frac{1}{2} (p_{0\uparrow}^2 + p_{0\downarrow}^2) \sum_{\mathbf{k}} \left[\langle T_\tau \tilde{p}_{\alpha\bar{\mathbf{k}}\downarrow}^\dagger(\tau) \tilde{p}_{\alpha\bar{\mathbf{k}}\downarrow}(0) \rangle + \langle T_\tau \tilde{p}_{\alpha\mathbf{k}\uparrow}^\dagger(\tau) \tilde{p}_{\alpha\mathbf{k}\uparrow}(0) \rangle \right] = 0, \quad (\text{S48})$$

to leading order in the boson correlations. As a result, the Green's function of the \tilde{p}_σ boson in $\Lambda_{\alpha\sigma}^{\text{incoh}}$ given in Eq. (S43) can be expressed in terms of those of the \tilde{d} - \tilde{e} bosons; leading to

$$\Lambda_{\alpha\sigma}^{\text{incoh}}(\tau) = \sum_{\mathbf{q}} \left\{ \rho_{\alpha\sigma}^e \langle T_\tau \tilde{e}_{\alpha\bar{\mathbf{q}}}^\dagger(\tau) \tilde{e}_{\alpha\bar{\mathbf{q}}}(0) \rangle + \rho_{\alpha\sigma}^d \langle T_\tau \tilde{d}_{\alpha\mathbf{q}}(\tau) \tilde{d}_{\alpha\mathbf{q}}^\dagger(0) \rangle + \rho_{\alpha\sigma}^{de} [\langle T_\tau \tilde{e}_{\alpha\bar{\mathbf{q}}}^\dagger(\tau) \tilde{d}_{\alpha\mathbf{q}}^\dagger(0) \rangle + \langle T_\tau \tilde{d}_{\alpha\mathbf{q}}(\tau) \tilde{e}_{\alpha\bar{\mathbf{q}}}(0) \rangle] \right\}, \quad (\text{S49})$$

where

$$\rho_{\alpha\sigma}^e = p_{\alpha 0\sigma}^2 - \tilde{p}_0^2, \quad \rho_{\alpha\sigma}^d = p_{\alpha 0\bar{\sigma}}^2 - \tilde{p}_0^2, \quad \rho_{\alpha\sigma}^{de} = p_{0\uparrow} p_{0\downarrow} - \tilde{p}_0^2; \quad \text{with } \tilde{p}_0^2 = 2d_0^4 / (p_{0\uparrow}^2 + p_{0\downarrow}^2). \quad (\text{S50})$$

We have verified that the above results hold even in the presence of pairing between the p_σ bosons. The normal and anomalous Green's function of the fluctuating doublons and holons in Eq. (S49) are readily evaluated using the boson wavefunctions obtained from the Bogoliubov transformation in Eqs. (S22),

$$\langle T_\tau \tilde{e}_{\alpha\bar{\mathbf{k}}}^\dagger(\tau) \tilde{e}_{\alpha\bar{\mathbf{k}}}(0) \rangle = \sum_{s=\pm} |u_{\alpha\mathbf{k}s}|^2 n_B[E_s^d(\mathbf{k})] e^{\tau E_s^d(\mathbf{k})} + |v_{\alpha\mathbf{k}s}|^2 \{1 + n_B[E_s^d(\mathbf{k})]\} e^{-\tau E_s^d(\mathbf{k})}, \quad (\text{S51})$$

$$\langle T_\tau \tilde{d}_{\alpha\mathbf{k}}(\tau) \tilde{d}_{\alpha\mathbf{k}}^\dagger(0) \rangle = \sum_{s=\pm} |v_{\alpha\mathbf{k}s}|^2 n_B[E_s^d(\mathbf{k})] e^{\tau E_s^d(\mathbf{k})} + |u_{\alpha\mathbf{k}s}|^2 \{1 + n_B[E_s^d(\mathbf{k})]\} e^{-\tau E_s^d(\mathbf{k})}, \quad (\text{S52})$$

$$\langle T_\tau \tilde{e}_{\alpha\bar{\mathbf{k}}}^\dagger(\tau) \tilde{d}_{\alpha\mathbf{k}}^\dagger(0) \rangle = \sum_{s=\pm} u_{\alpha\mathbf{k}s} v_{\alpha\mathbf{k}s}^* n_B[E_s^d(\mathbf{k})] e^{\tau E_s^d(\mathbf{k})} + u_{\alpha\mathbf{k}s}^* v_{\alpha\mathbf{k}s} \{1 + n_B[E_s^d(\mathbf{k})]\} e^{-\tau E_s^d(\mathbf{k})}, \quad (\text{S53})$$

$$\langle T_\tau \tilde{d}_{\alpha\mathbf{k}}(\tau) \tilde{e}_{\alpha\bar{\mathbf{k}}}(0) \rangle = \sum_{s=\pm} u_{\alpha\mathbf{k}s}^* v_{\alpha\mathbf{k}s} n_B[E_s^d(\mathbf{k})] e^{\tau E_s^d(\mathbf{k})} + u_{\alpha\mathbf{k}s} v_{\alpha\mathbf{k}s}^* \{1 + n_B[E_s^d(\mathbf{k})]\} e^{-\tau E_s^d(\mathbf{k})}, \quad (\text{S54})$$

where the $n_B(E) = 1/[\exp(\beta E) - 1]$ is the boson distribution function. Substituting Eqs (S51-S54) in Eq. (S49), inserting the result into Eq. (S45), and carrying out the imaginary-time integral, we obtain the final expression for the incoherent part of the ISF,

$$N_\alpha^{\text{incoh}}(\omega) = \pi \sum_{\mathbf{k}, \mathbf{q}, \sigma} \sum_{s=\pm} r_{\alpha\sigma}^2 \left\{ \left[\rho_{\alpha\sigma}^e |u_{\alpha\mathbf{q}s}|^2 + \rho_{\alpha\sigma}^d |v_{\alpha\mathbf{q}s}|^2 + 2\rho_{\alpha\sigma}^{de} \text{Re}(u_{\alpha\mathbf{q}s}^* v_{\alpha\mathbf{q}s}) \right] |\tilde{u}_{\alpha\mathbf{k}\sigma}|^2 \delta[\omega + E_{\mathbf{k}}^f + E_s^d(\mathbf{q})] \right. \\ \left. + \left[\rho_{\alpha\sigma}^e |v_{\alpha\mathbf{q}s}|^2 + \rho_{\alpha\sigma}^d |u_{\alpha\mathbf{q}s}|^2 + 2\rho_{\alpha\sigma}^{de} \text{Re}(v_{\alpha\mathbf{q}s}^* u_{\alpha\mathbf{q}s}) \right] |\tilde{v}_{\alpha\mathbf{k}\sigma}|^2 \delta[\omega - E_{\mathbf{k}}^f - E_s^d(\mathbf{q})] \right\}. \quad (\text{S55})$$

Remarkably, at $U = 0$, where $p_{0\sigma} = d_0 = e_0 = 1/2$, $\rho_{\alpha\sigma}^e = \rho_{\alpha\sigma}^d = \rho_{\alpha\sigma}^{de} = \rho_{\alpha\sigma}^{ed} = 0$ in Eq. (S55) and the incoherent spectral function is completely suppressed, recovering the noninteracting limit. At any finite U , our numerical calculations show that the spectral sum rule is approximately satisfied and an overall normalization factor close to one is applied to Eq. (S55) to ensure $W_{\alpha\sigma}^{\text{incoh}} + W_{\alpha\sigma}^{\text{coh}} = 1$ in the plotted ISF figures. The obtained ISF has been shown in Fig. 3 of the main text. To illustrate in more detail the transfer of the coherent QP weight to the incoherent part of the spectral function with increasing Hubbard U , we plot in Fig. S4 separately the coherent (shaded red areas), incoherent (blue solid lines), and the total (black solid lines) ISF for several values of the Hubbard U across the Mott transition.

-
- [1] G. Kotliar and A. E. Ruckenstein, Phys. Rev. Lett. **57**, 1362 (1986).
 - [2] R. Frésard and K. Doll, in *The Hubbard Model: Its Physics and Mathematical Physics*, Editors: D. Baeriswyl, et. al., page 385-392, Plenum Press (1995).
 - [3] G.D. Mahan, Many-particle physics, 2nd Ed. (Plenum Press, New York).
 - [4] Z. Wang, Y. Bang, G. Kotliar, Phys. Rev. Lett. **67** 2733 (1991). Z. Wang, International Journal of Modern Physics B**6**, 155 (1992).